

MEASURING THE 3D SHAPE OF X-RAY CLUSTERS

Johan Samsing¹, Andreas Skielboe¹, Steen H. Hansen¹

¹ *Dark Cosmology Centre, Niels Bohr Institute, University of Copenhagen, Juliane Maries Vej 30, 2100 Copenhagen, Denmark*

ABSTRACT

Observations and numerical simulations of galaxy clusters strongly indicate that the hot intracluster x-ray emitting gas is not spherically symmetric. In many earlier studies spherical symmetry has been assumed partly because of limited data quality, however new deep observations and instrumental designs will make it possible to go beyond that assumption. Measuring the temperature and density profiles are of interest when observing the x-ray gas, however the spatial shape of the gas itself also carries very useful information. For example, it is believed that the x-ray gas shape in the inner parts of galaxy clusters is greatly affected by feedback mechanisms, cooling and rotation, and measuring this shape can therefore indirectly provide information on these mechanisms. In this paper we present a novel method to measure the three-dimensional shape of the intracluster x-ray emitting gas. We can measure the shape from the x-ray observations only, i.e. the method does not require combination with independent measurements of e.g. the cluster mass or density profile. This is possible when one uses the full spectral information contained in the observed spectra. We demonstrate the method by measuring radial dependent shapes along the line of sight for **CHANDRA** mock data. We find that at least 10^6 photons are required to get a $5 - \sigma$ detection of shape for an x-ray gas having realistic features such as a cool core and a double powerlaw for the density profile. We illustrate how Bayes' theorem is used to find the best fitting model of the x-ray gas, an analysis that is very important in a real observational scenario where the true spatial shape is unknown. Not including a shape in the fit may propagate to a mass bias if the x-ray is used to estimate the total cluster mass. We discuss this mass bias for a class of spacial shapes.

1. INTRODUCTION

Galaxy clusters are the largest bound objects in the universe and they provide unique and independent information on the cosmological evolution. The standard LCDM parameters and a possible redshift varying dark energy component has accurately been measured and constrained from cluster observations in a variety of ways (Vikhlinin et al. (2009b), Allen et al. (2011), Mantz et al. (2010), Allen et al. (2008), Vikhlinin et al. (2009a)). They reveal the distant universe behind them through gravitational magnification (Kneib et al. (2004), Amanullah et al. (2011), Bradley et al. (2011)), and they are even sensitive to the initial perturbations of our universe (Fedeli et al. (2009), Chongchitnan & Silk (2011), Sartoris et al. (2010)). Clusters not only serve as excellent laboratories for constraining the standard cosmology, but because of their relative high mass and cosmological size they also provide a unique possibility to test general relativity itself in several independent ways, e.g. from measurements of cosmic growth (Rapetti et al. (2010)) to gravitational redshift (Wojtak et al. (2011)) and gravitational waves (Yoo et al. (2009)). Other probes have also been suggested such as lensing, cluster abundance and

the integrated Sachs-Wolfe effect (Jain & Zhang (2008)). Despite their importance in modern cosmology, basic properties such as spatial shape is still not well measured for individual clusters. One reason is simply that the main part of a cluster is composed of dark matter which can only be measured indirectly by its gravitational interaction. The indirect measurements of the dark matter and its radial distribution are usually done using either lensing (Postman et al. (2011), Stark et al. (2007)), by studying the dynamics of the intracluster galaxies (Wojtak & Lokas (2010), Lokas & Mamon (2003), Lemze et al. (2009)) or by the hot baryonic x-ray emitting gas located in the inner regions of all clusters (for a review of x-ray physics and applications see e.g. Sarazin (1988)). Especially observations of the intracluster x-ray gas in terms of spacial shape, density and temperature profiles, play a key role for estimating local properties of the cluster. Many earlier studies assume a spherical shape of the gas (Pointecouteau et al. (2005), Host & Hansen (2011), Kaastra et al. (2004), Piffaretti et al. (2005), Hansen & Piffaretti (2007), Rapetti et al. (2010)), however there are several strong motivations why a precise estimation of the shape is interesting. One is a precise estimation of the cluster mass profile. This profile can directly be

measured if the radial shape, temperature and density profiles of the gas are known and the gas is in hydrostatic equilibrium. Only recently it was shown that allowing the gas to have a triaxial shape is necessary for the estimated mass profile from x-ray to agree with the mass estimated from lensing (Morandi et al. (2010), Sereno & Umetsu (2011), Morandi et al. (2011)), a result in good agreement with numerical simulations (Hayashi et al. (2007)). This overall triaxiality is mostly due to the underlying shape of the dark matter potential. However, in the central cluster regions it is believed that a possible non-spherical x-ray shape is more affected by microphysical processes such as radiative cooling, turbulence and different feedback mechanisms (Lau et al. (2011)) than the dark matter potential shape is. These mechanisms change the gas shape into having relatively high ellipticity towards the center compared to the underlying dark matter potential shape. It is therefore possible to infer properties of these mechanisms if the shape of the gas, temperature and density profiles are known to high precision.

In this paper we suggest and develop a method from which a possible radial dependent shape of an x-ray gas can be extracted from the x-ray observations only. We explicitly demonstrate the possibilities for measuring the shape by fitting to *CHANDRA* mock data and we estimate the mass bias if a shape is not treated correctly in the fitting. The method we use is a parametrized approach, i.e. we assume that the shape and profiles can be described by a set of well defined functional forms. We also discuss the complications of choosing the best set of functions, i.e. a model, to describe the data.

The paper is organized in the following way; The method for measuring shape is explained in section 2.1. We apply the method in section 3 on *CHANDRA* mock data. We discuss how to quantify the goodness of fit in section 3.3. Mass bias from not including the shape in the fitting is discussed in section 4.

2. EXTRACTING 3D X-RAY INFORMATION FROM 2D OBSERVATIONS

An intracluster x-ray emitting gas has a three dimensional extension, spherical or not, but an observer will only see the two dimensional projected image on the sky. Therefore, a given observed spectrum is a sum of all emission spectra along the line of sight through the gas (for a discussion see figure 1).

Each spectrum has a spectral shape determined by the local temperature and a scaling proportional to the local density squared (Sarazin (1988)). Mathematically, no unique mapping can construct the true three dimensional shape, density and temperature profiles using only the observed two dimensional image. However, if one makes prior assumptions it can be done. For instance by assuming that the gas is spherical the density and temperature profiles can be found. From this assumption

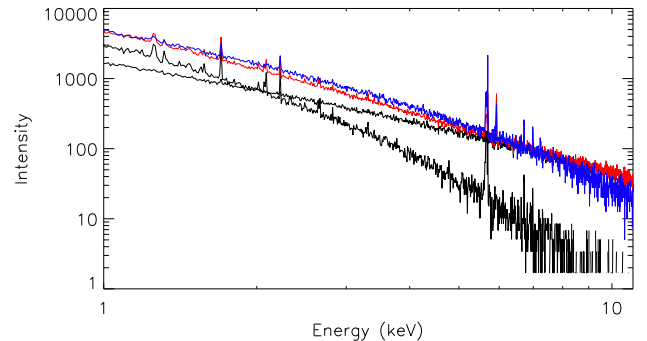


Fig. 1.— Simplified description of how information propagates from three dimensions to a two dimensional observation. The figure shows two free-free emission X-ray spectra (the two lowest spectra at 3 keV in black), the sum of these two spectra (second spectrum from the top at 3 keV in red), and a best fit free-free spectrum to the red spectrum (upper spectrum at 3 keV in blue). This could correspond to a part of a gas with two temperature components (one spectrum for each component) projected along the line of sight (observed data). By comparing the blue and the red spectra one notices that the red spectrum cannot simply be fitted accurately with a free-free emission X-ray spectrum. This is true in the general case; a sum of free-free spectra cannot in general be fitted by another free-free spectrum. If the observed red spectrum is not correctly fitted one could incorrectly conclude the presence of either a *non thermal hard X-ray excess* or *non thermal soft X-ray excess* component, depending on the shape of the red spectrum. In our example one would report a *hard excess* component since the red spectrum is above the blue in the tail (see e.g. Fusco-Femiano et al. (2005) for a hard excess discussion for Abell 2256).

several previous groups have measured the temperature and density profiles of the x-ray gas using either projection or de-projection techniques (see e.g. the *XSPEC* packages 'deproject' and 'project')

The method we present in this paper for extracting three-dimensional information relies on the assumption that the x-ray gas shape, density and temperature profiles can be described by parametrizations. This means that shape and profiles are believed to be well described by a set of functions. In contrast to several previous studies we use the whole spectral information from the integrated observed picture of the x-ray gas. It means that we take into account that the actual observed spectra is a sum of spectra along the line of sight, and can not simply be fitted by a single free-free spectrum. See figure 1 for a discussion. We allow a radial dependent shape in contrast to previous studies. The tradeoff for including this extra freedom is that we limit our analysis to structures that are seen spherical on the sky. This is for purely practical reasons: in theory the fitting method we describe is not limited by this assumption, but with present day available data it is simply not possible to

resolve a radial dependent shape if the 3D-shape and orientation of the gas is completely free to vary. In other words, the symmetry in the sky makes it possible to extract higher order corrections to the usual assumption about either sphericity or triaxiality with constant axis ratios. The method and procedure will be described in the following sections, and technical details are found in the appendix together with illustrations of generated spectra and an x-ray structure.

2.1. Fitting shape and profiles using the parameterization approach

The procedure needed in order to measure spatial shape, temperature and density profiles of an observed x-ray gas using the parametrization approach is as following: First we choose a model, i.e. a set of parameterizations, that are believed to generally describe the form of density, temperature and spatial shape (along the line of sight) for the observed structure. The chosen parameterizations must be sufficiently general to accurately describe observations of real and simulated structures. We then calculate the agreement between an artificial generated dataset (see appendix section A.0.1 for how we generate artificial datasets and mock data) created from the chosen model given a specific combination of parameter values and the observed dataset. In our case we quantify the agreement by a simple χ^2 statistic which simply can be related to a probability by $\exp(-\chi^2/2)$ when the noise is gaussian. This routine of comparing artificial generated datasets with the observed dataset is then repeated for a wide range of parameter value combinations until a good estimate of the underlying probability distribution function (PDF) for our model has been made. For this we use standard Monte Carlo techniques as described in section A.0.2 in the appendix. From the parameter combination having the maximum PDF value, the best estimate for profiles and shape, given our prior input parameterizations, can then be made. The overall procedure can then be repeated for different models, until the best model is found. We will discuss this in more detail in section 3.3.

3. RESULTS FROM FITTING SHAPE AND PROFILES OF SELECTED X-RAY MODELS

In the following we show the possibilities of measuring radial profiles of non-spherical x-ray structures with varying radial dependent shape along the line of sight. As briefly discussed in the end of section 2, we only consider structures that are spherical on the sky. We consider two simulated structures in our analysis; First a simple toy model to clearly illustrate the method, and second a more realistic model with features such as a cool core and a double powerlaw for the density profile. The shape parameterizations are described later. In this part of the analysis we fit for profiles and shape using the same set

of parameterizations that are used to generate the data. In this way we get the cleanest picture of how a shape signal propagates to observables.

We present results in terms of a virial radius r_v . The shape, temperature and density profiles we use, are consistent with a virial radius similar to r_{500} (Vikhlinin et al. (2006)).

3.1. A simple toy model

We consider a dataset denoted by 'shM1' where the density and temperature profiles are modeled by simple broken powerlaws

$$\rho(r) = n_0(1 + (r/r_c)^2)^{-3\beta/2} \quad (1)$$

$$T(r) = T_0(1 + (r/r_t)^2)^{-b} \quad (2)$$

known as beta-models. Parameter n_0 acts as a normalization factor and is regulated such that the artificial dataset has a fixed number of total (photon) counts. The shape parametrization we consider is a simple linear function for ellipticity

$$\epsilon_1(r) = s_2 \cdot r + s_1 \quad (3)$$

where $\epsilon \equiv b/a$ is defined as the ratio between the radius perpendicular to the observer (b) and the radius along the line of sight (a) of the observer. The parameter values for shM1 are listed in table 1, and figure 2 shows the corresponding shape and profiles. The chosen parameters for the density and temperature profiles are in fair agreement with typical observed values. The priors on the shape parametrization we use in this example are: a) $0.2 < b/a \leq 1$ and b) $\epsilon_1(r/r_v = 1) > 0.5$. In general, a structure could naturally have an axis ratio $b/a \geq 1$ and still be spherical on the sky, and therefore in a scenario where no prior shape information is available, shapes with $b/a \geq 1$ must be included in the fit as well.

The left plot of figure 3 shows the maximized PDFs for the fitted density, temperature and shape parameters for a total of $3 \cdot 10^4$ photon counts (≈ 10 ks CHANDRA observation of A1689). The width of the projected PDFs, i.e. a measure of the fitting error for each parameter, is simply related to the number counts by $\sim 1/\sqrt{N}$ where N is the number of photons. The right plot of figure 3 shows the corresponding correlation matrix defined in the usual way as $CORR(X, Y) = \frac{\langle (X - \mu_X)(Y - \mu_Y) \rangle}{\sigma_X \sigma_Y}$ where X, Y are random variables with expectation values μ_X, μ_Y and standard deviations σ_X, σ_Y . In our case, to find e.g. the correlation coefficient $CORR(p_i, p_j)$ X must be replaced with the vector of MCMC sampled p_i values and Y of sampled p_j values. The correlation matrix is symmetric by construction and the shading goes from 0 (black) to 1 (white). The correlation matrix can be divided up in several regions. On the plot is highlighted a region bounded

TABLE 1
SIMULATED PROFILES

Model	equation	n_0	r_c	β	$\alpha/2$	T_0	a	b	r_t	s_1	s_2
shM1	1,2,3	...	0.11	0.6	0	5.0	0	0.14	0.09	0.3	0.83
shM2	4,6,7	...	0.15	0.76	1.2	4.3	2.45	0.7	0.13	0.94	0.2

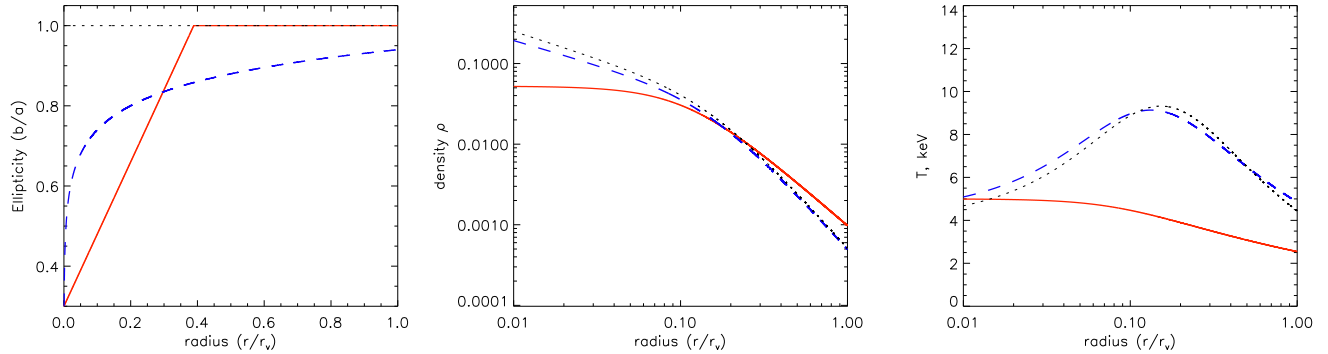


Fig. 2.— Shape and profiles for dataset shM1 (red, solid lines) and shM2 (blue, dashed lines). Left: Ellipticity of the x-ray gas along the line of sight. Central: x-ray gas density profile. Right: x-ray gas temperature profile. Black, dotted lines: Best estimate from a MCMC fitting to shM2 using the true parametrizations for temperature and density, but assuming spherical symmetry.

by a dotted line and a solid line. The dotted region is the part which shows the correlation between temperature and density and the solid is the region that shows the correlation between temperature and shape. In the lower left corner of the correlation matrix is the region showing the correlation between shape and density. In general we see that the temperature is weakly correlated with the rest of the parameters, especially when compared to the correlation between shape and density. The physical reason is simply that their individual contributions to a spectrum by nature are completely different; temperature affects the spectral form, but density and shape affect only the normalization. This is clearly seen in the analytic form for the bremsstrahlung spectrum $I(T, \rho, \nu) \propto \rho^2 T^{-1/2} \exp(-\nu/T)$ (Sarazin (1988)).

Among our chosen priors, the prior on s_1 ($s_1 > 0.2$) is the one that affects the shape of the PDFs the most. Besides a trivial truncation on the s_1 parameter axis it is also responsible for especially the truncation (or skewness) of the n_0 distribution. The reason is the relative strong correlation between these two parameters. This correlation is clearly seen on the correlation matrix and can be understood in the following way: The degree of constant ellipticity captured by s_1 effectively acts as mass scaling term when the structure is projected along the line of sight. This is simply because an ellipticity “stretches” the structure and therefore “allows” more mass along the line of sight. This is exactly how n_0 affects the projected dataset too. So if we increase the overall

scaling (increasing n_0) we can compensate by decreasing the ellipticity (increasing s_1), that means the lower truncation of s_1 also shows up as a lower truncation on n_0 . In fact, a constant ellipticity along the line of sight ϵ is completely degenerate with the overall density scaling ρ_0 by $\rho_0^2 \epsilon$. This is an intrinsic degeneracy and can only be broken by including other observations, e.g. SZ observations which effectively traces $\rho_0 \epsilon T$ (see e.g. Planck Collaboration et al. (2011), De Filippis et al. (2005), Conte et al. (2011), Sereno et al. (2011)).

The overall conclusion from the fitting results is that the parameter values specifying the true shape as well as temperature and density are exactly reconstructed. This is an ideal case, but it is clearly showing that temperature, density and shape in principle can be separated.

From the correlations we can conclude that the temperature profile is well and almost independently fitted. In perspective of optimizing the fit for shape, this also implies that independent measurements of the density will directly result in a better fit for the shape.

3.2. A more realistic model

We now perform an analysis on a dataset, denoted by ‘shM2’, describing a structure with cool core and a double powerlaw for the density profile. Including these features are motivated by real observations (Vikhlinin et al. (2006)). The temperature and density profiles are now parameterized as

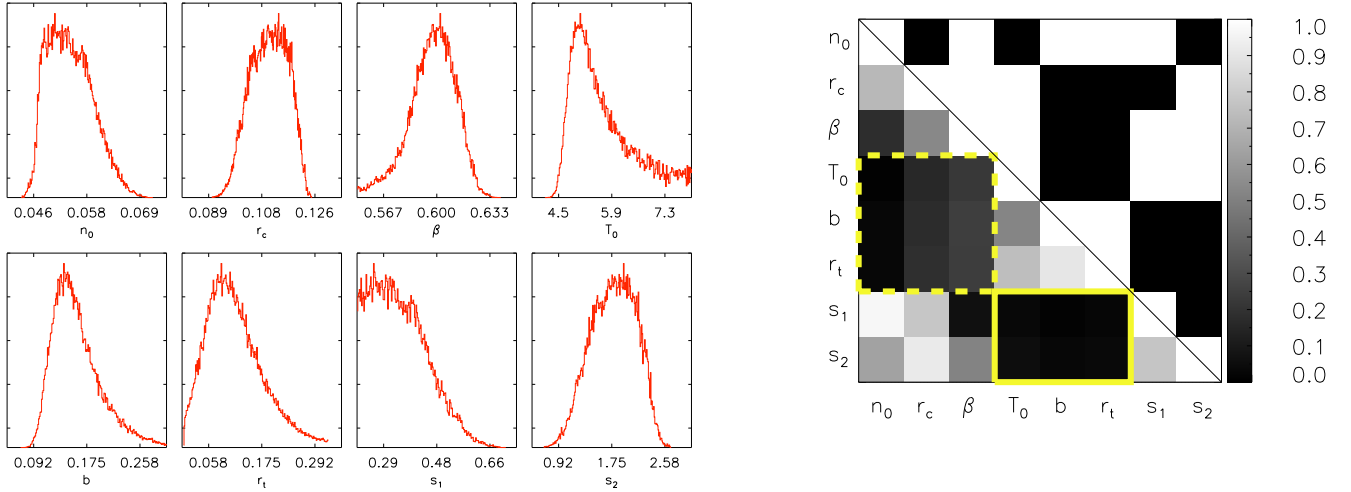


Fig. 3.— Fitting results for dataset 'shM1'. Left: Maximized PDFs along each of the 8 parameters in the model. Right: Correlation matrix for the 8 parameters in the model. The lower left part of the matrix shows the absolute value of the correlation coefficient where the upper right corner shows the sign of the coefficient in black (negative) and white (positive). The results are based on a dataset scaled to have a total of $3 \cdot 10^4$ photon counts.

$$\rho(r) = n_0 \frac{(r/r_c)^{-\alpha/2}}{(1 + (r/r_c)^2)^{3\beta/2 - \alpha/4}} \quad (4)$$

$$T(r) = T_0 \frac{1 + a(r/r_t)}{(1 + (r/r_t)^2)^b}, \quad (5)$$

and the shape is parameterized by

$$\epsilon_2(r) = s_2 \cdot \log_{10}(r) + s_1 \quad (6)$$

This shape parametrization approximately describes the gas shape seen in the inner parts ($r \leq r_{500}$) of clusters in numerical simulations (Lau et al. (2011)). We use the same shape priors as used in the previous toy model example. A list of temperature and density parameterizations are found in (Vikhlinin et al. (2006)).

The true parameter values for 'shM2' are listed in table 1 and the corresponding shape and profiles are plotted in figure 2. Figure 4 shows the PDF and the correlation matrix for the 10 parameter model fitting.

An inner density slope captured by α is now one of the new parameters compared to the toy model. Since both the shape and this inner slope have a logarithmic dependence, there is a strong correlation between α and s_2 . This is clearly seen in the correlation matrix and the PDF plot where the lower cut on s_2 directly relates to the skewness in the α distribution. This freedom in the inner slope is the main reason for the fitting to require many more photons than the toy model. This is discussed in more detail in section 3.3 below.

The overall conclusion is that the true parameter values are reconstructed, but to keep down the statistical errors a relative high number of photons are required. This is mostly due to the similar parameterizations for shape and density. In agreement with intuition, we see that it is much harder to extract a logarithmic shape when the density is varying logarithmically too, compared to e.g. a linear dependent shape. From the correlation matrix we see that the temperature fitting is nearly unaffected as we also concluded in the previous toy model example.

3.3. Quantifying the goodness of fit

In this section we will discuss how to quantify the goodness of fit for the parameters within a given model, as well as the goodness of fit for the model itself relative to other competitive models.

3.3.1. Individual parameters within a model

The best fit parameter values for a given model are located at the likelihood maximum, or the minimum χ^2 if the measurement noise is gaussian. To quantify the goodness of the fit is not unique in the same way. To quantify this one must often combine statistical estimators with prior knowledge. An often used estimator is the reduced chi square, $\chi_{red}^2 = \chi^2/K$, where K is the number of degrees of freedom. However, this estimator has two major problems. First, χ^2 itself have a significant noise due to random noise of the data, and second, the number of degrees of freedom is not in general well defined (Andrae et al. (2010)). Another, maybe more intuitive, estimator

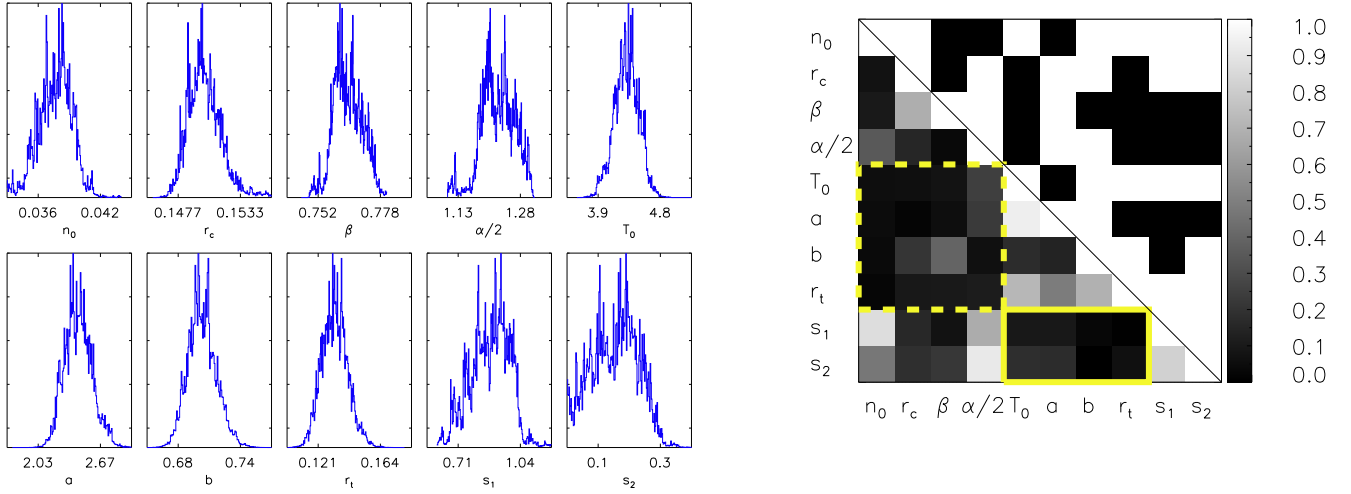


Fig. 4.— Fitting results for dataset ‘shM2’. Left: Maximized PDFs along each of the 10 parameters in the model. Right: Correlation matrix for the 10 parameters in the model. The lower left part of the matrix shows the absolute value of the correlation coefficient where the upper right corner shows the sign of the coefficient in black (negative) and white (positive). The results are based on a dataset scaled to have a total of $1 \cdot 10^6$ photon counts.

is the the ratio \hat{p}_i/σ_i where \hat{p}_i is the best estimate for parameter p_i and σ_i the associated standard deviation. If we denote this ratio by n we can quantify the goodness of fit by reporting n for each parameter or the minimum n for the whole model. For the fitting examples we presented above, it is then of interest to know the number of photons required for e.g. a minimum $n = 5$ (or $5 - \sigma$) detection for all parameters. We will investigate this in the following. Figure 5 shows the ratio \hat{p}_i/σ_i as a function of total photon counts for the parameter that have the largest ratio, i.e. the parameter which is most difficult to estimate, for the two structures ‘shM1’ (left plot) and ‘shM2’ (right plot). In the ‘shM1’ example the most difficult parameter to estimate in terms of n is r_t , to reach a minimum $5 - \sigma$ detection of this (and thereby for each parameter of the whole model) we find from the figure that more than $\approx 2.5 \cdot 10^5$ photons are required. If we instead only require that the shape parameters must be estimated with a minimum $5 - \sigma$ each we find a limit of $\approx 1.4 \cdot 10^5$ photons, or roughly a factor of 2 less compared to an overall $5 - \sigma$ detection. Following the same procedure for the more realistic example ‘shM2’ we find that a minimum of $\approx 4.2 \cdot 10^6$ photons are required for a minimum $5 - \sigma$ detection on all parameters. The same number of photons are required for the shape fitting because s_2 is the most difficult parameter to estimate in terms of n .

3.3.2. Model comparison

Assuming that the quantities we try to measure for a gas can be parameterized, we still have the problem that we have no idea of how the “true” or best parametrization for the gas looks like in a real observation. This means e.g. that a set of shape parameters defined in a specific gas parametrization do not have to describe a real shape at all. The parameters could in principle just capture higher order corrections to the density profile, because of the general tight correlation between shape and density. In this case, the real problem is to realize that your model does not return information about the system in the way you believe. The question is therefore how to quantify how a specific model performs relative to one or several other competitive models. A useful measure of this can be found using Bayes’ theorem. From this theorem it is possible to calculate the relative probability, also known as the *posterior odds*, of two competing models (Jenkins & Peacock (2011), Trotta (2008)). In the case where we assume flat parameter and model priors, the posterior odds ratio reduces to the simple ratio

$$\mathcal{F}(H_1, H_0) = \int \mathcal{L}(D | H_1, \beta) d\beta / \int \mathcal{L}(D | H_0, \alpha) d\alpha \quad (7)$$

where $\mathcal{L}(D | H, \beta)$ is the likelihood for getting the data D given the model H which depends on the parameter-set β . This ratio \mathcal{F} is often denoted the *evidence ratio* between model H_1 and H_0 . Model H_0 is often a ‘null’ or default model where H_1 is a competing and often more

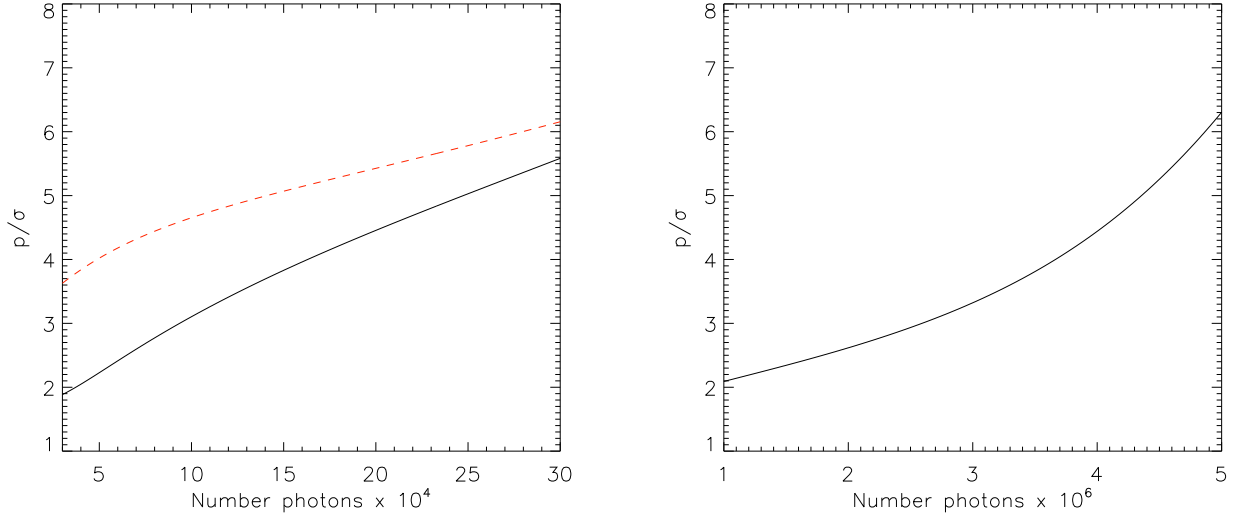


Fig. 5.— The plot shows the ratio $n = \hat{p}_i/\sigma_i$ for the most difficult parameters to estimate when fitting for structure 'shM1' (left figure) and 'shM2' (right figure) in black lines. The worst estimated parameter in terms of n when fitting to 'shM1' is r_t and s_2 when fitting to 'shM2'. The red dashed line in the left plot shows the ratio for parameter s_1 which is the worst determined of the two shape parameters when fitting to 'shM1'. It can be read of the figure that $\approx 4.2 \cdot 10^6$ number total photon counts are required for a minimum $5 - \sigma$ detection on all parameters when fitting to 'shM2'. The radial and spectral binning is kept constant in the plot.

complicated model. In our case, H_0 could be a model assuming spherical symmetry and H_1 a model allowing the shape to vary. The evidence threshold, or critical threshold, between rejecting or accepting a competitive model is often taken to be Jeffreys threshold 1:148 (Jenkins & Peacock (2011)). Let us now go through a few examples.

First suppose we want to compare two models, M_1 and M_2 , given the data set shM2. Both models are using the correct parameterizations for temperature and density, but not the same parametrization for shape; model M_1 includes the true parametrization of shape in the fitting, but model M_2 assumes spherical symmetry. We can now use Bayes' theorem to show if e.g. a $5 \cdot 10^5$ photon exposure carries enough information to distinguish between M_1 and M_2 . Performing the two integrals in equation 7 for a $5 \cdot 10^5$ photon exposure we find $\mathcal{F} \sim 900$, i.e. we can correctly conclude that M_1 is strongly favored over M_2 . The slightly biased estimations for the density and temperature when shM2 is fitted assuming M_2 is seen in figure 2.

Another scenario could be that we fit the shape with a parametrization that is different from the true one. In that case, suppose we fit dataset shM1 with two models M_1 and M_2 . Both of them are using the true temperature and density parameterizations, but model M_1 is using equation 3 for the shape parameterization in contrast to model M_2 that is using equation 6. For a $3 \cdot 10^4$ photon exposure we find $\mathcal{F} \sim 6700$, concluding correctly that M_1 is strongly favored over M_2 .

The last example is a case where the true structure has temperature and density profiles as 'shM2', but have a spherical shape. We now make a fit including shape, but we use equation 1, i.e. a simple beta-model, to describe the density instead of the true equation 4 that has one extra degree of freedom. The interesting thing is now that the best fit using the beta-model will show clear detection of shape away from spherical. This is seen on figure 6. The under fitted density profile is simply compensated by allowing a non-spherical shape in the inner parts. This is a false detection. In a real case where the true shape of the gas is not known, this can be very hard to realize. Comparing this fit using Bayes' theorem with a fit using the more general density profile in equation 4 we find $\mathcal{F} \sim 300$ for a $6 \cdot 10^4$ photon exposure. Which correctly means a spherical model is favored.

It is possible to write up a simple scaling relation between number photons and the evidence ratio given that the PDF approximately can be described by a multidimensional gaussian near its peak; Assume from a N_2 photon exposure we have calculated the evidence ratio \mathcal{F}_{N_2} between two models M_A and M_B , from that we can simply calculate the ratio \mathcal{F}_{N_1} for a N_1 photon exposure by $\mathcal{F}_{N_1} \approx \mathcal{F}_{N_2} (P_{M_A}^*/P_{M_B}^*)^{(N_1/N_2-1)}$ where P^* is the value of the PDF at its maximum for the N_2 photon exposure. Here we have used the analytical solution to equation 7 (see e.g. Jenkins & Peacock (2011) eq. 8). This scaling relation can be useful for forecasting the case where a correct integration is limited by, e.g. computational power. However, this estimator can be relative noisy because of

its dependence on the value at the PDF maximum. One way to reduce this scatter could be to fit a gaussian to the PDF near its peak.

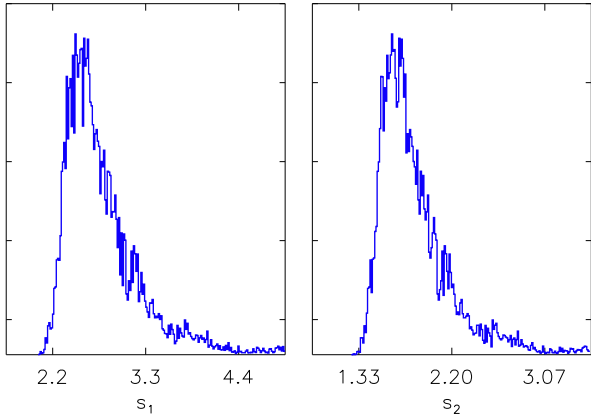


Fig. 6.— Example of a false detection of shape. Maximized PDFs for the two shape parameters s_1 and s_2 when fitting to a spherical version of 'shM2', but using a beta-model for the density instead of the true equation 4. The peaks in the shape parameters are real, but are not representing shape. They reflect that the parameterization for density used in the fitting is not having enough freedom to describe the variation of density in the inner parts. As described in the text, Bayes' theorem can be used to quantify if this is a true signal of shape or not. The plot is for a $6 \cdot 10^4$ photon exposure.

4. X-RAY GAS SHAPE AND CLUSTER MASS BIAS

By knowing the 3D x-ray gas temperature and density profiles one can calculate the underlying total cluster density, and hence mass, by combining the hydrostatic equilibrium (HE) equation

$$\nabla(\rho_{gas}T_{gas}) = -\rho_{gas}\nabla\Phi_{total} \quad (8)$$

with the poisson equation

$$\nabla^2\Phi = 4\pi G\rho_{total} \quad (9)$$

where the index 'total' indicates that the contribution is from both gas and dark matter. When x-ray observations are possible for a cluster and the x-ray gas is in HE, this method is one of the most precise ways to estimate the cluster mass as a function of radius within the visible x-ray region. However, as we can see, the estimated cluster mass will be wrong if the gas is not in HE or if ρ_{gas} , T_{gas} is not correctly known. One way of misestimating ρ_{gas} and T_{gas} is fitting a spherical model to data for an intrinsic non-spherical gas structure. Depending on the shape, this assumption will propagate to a bias in the estimated total cluster mass. In this section we will study

the cluster mass bias as a function of different shapes along the line of sight.

4.1. Mass Bias

The upper plot in figure 7 shows the shape along the line of sight for four different x-ray structures. We take the four structures to have temperature and density profiles similar to shM2, but different spacial shapes. Fitting temperature and density profiles to these four structures assuming spherical symmetry, will result in biased mass profiles. The ratio between the biased and the true mass profile is shown in the lower plot in figure 7. We have only included the mass contribution within r_v . Taking the rest of the mass of the cluster into account, requires an extrapolation of the dark matter potential form beyond the visible x-ray region. This is necessary when combining or comparing with other mass probes such as lensing.

As seen on the plot, the shapes we are considering leads to small biases at the 10 percent level, dependent on the radius. This difference can be important for doing future precision cosmology using clusters. However, at this level the degree of hydrostatic equilibrium may lead to higher uncertainties in the mass estimation (Lau et al. (2009), Piffaretti & Valdarnini (2008), Cavaliere et al. (2011)).

5. CONCLUSIONS

We have presented a new method for measuring a radial dependent shape along the line of sight of the intra-cluster x-ray emitting gas. The method uses the assumption that the shape, temperature and density profiles can be described by parameterized functions. Compared to several previous studies, we use the whole spectral information. Using this method we have demonstrated the possibilities for measuring shape on CHANDRA mock data.

We find that around 10^6 photons are required to get a $5 - \sigma$ detection of shape when fitting to a model showing realistic features of the gas, such as cool core and a double powerlaw for the density profile. We have seen, by presenting correlations matrices, that density and shape have a strong correlation, whereas temperature is essentially uncorrelated. This strong correlation indicates that independent measurements of the density profile can strongly improve the estimation of shape.

We demonstrated that Bayes' theorem very effectively can be used to compare different prior input models for our approach. This is of great importance since the actual science one extracts in the end has to be read off from the input model.

Finally we showed the effect on the mass profile estimation from assuming spherical symmetry when fitting structures with non-spherical shapes. Within our considered class of shapes, we found the mass estimation to be biased at the 10% level.

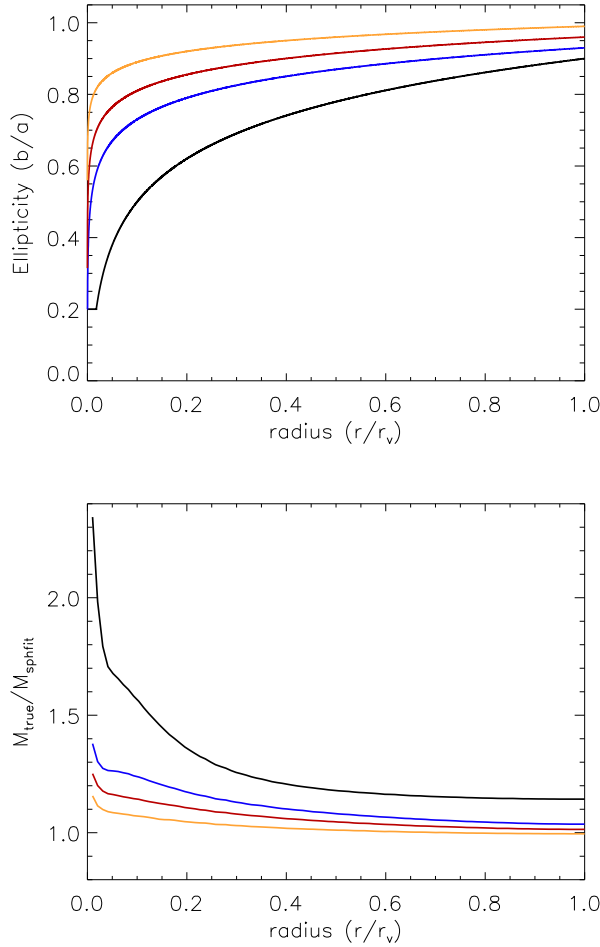


Fig. 7.— Mass bias from assuming spherical symmetry when fitting to non-spherical x-ray structures. Upper: Shape along the line of sight for four different prolate x-ray structures. Lower: Ratio between the true total mass profile and the mass profile estimated from a spherical fit to dataset 'shM2' having the shapes shown at the upper plot. Opposite bias (i.e. $M_{true}/M_{sphfit} < 1$) is expected for oblate structures.

In a future paper we will use our framework on real data.

We warmly thank Martina Zamboni for useful discussions. The Dark Cosmology Centre is funded by the Danish National Research Foundation.

REFERENCES

- Allen, S. W., Evard, A. E., & Mantz, A. B. 2011, *ARA&A*, 49, 409
- Allen, S. W., Rapetti, D. A., Schmidt, R. W., Ebeling, H., Morris, R. G., & Fabian, A. C. 2008, *MNRAS*, 383, 879
- Amanullah, R., et al. 2011, ArXiv e-prints
- Andrae, R., Schulze-Hartung, T., & Melchior, P. 2010, ArXiv e-prints
- Arnaud, K. A. 1996, in *Astronomical Society of the Pacific Conference Series*, Vol. 101, *Astronomical Data Analysis Software and Systems V*, ed. G. H. Jacoby & J. Barnes, 17–+
- Bradley, L. D., et al. 2011, ArXiv e-prints
- Cavaliere, A., Lapi, A., & Fusco-Femiano, R. 2011, *A&A*, 525, A110+
- Chib, S., & Greenberg, E. 1995, 49, 327
- Chongchitnan, S., & Silk, J. 2011, ArXiv e-prints
- Conte, A., de Petris, M., Comis, B., Lamagna, L., & Gregori, S. 2011, *A&A*, 532, A14+
- De Filippis, E., Sereno, M., Bautz, M. W., & Longo, G. 2005, *ApJ*, 625, 108
- Fedeli, C., Moscardini, L., & Matarrese, S. 2009, *MNRAS*, 397, 1125
- Fusco-Femiano, R., Landi, R., & Orlandini, M. 2005, *ApJ*, 624, L69
- Hansen, S. H., & Piffaretti, R. 2007, *A&A*, 476, L37
- Hayashi, E., Navarro, J. F., & Springel, V. 2007, *MNRAS*, 377, 50
- Host, O., & Hansen, S. H. 2011, *ApJ*, 736, 52
- Jain, B., & Zhang, P. 2008, *Phys. Rev. D*, 78, 063503
- Jenkins, C. R., & Peacock, J. A. 2011, *MNRAS*, 413, 2895
- Kaastra, J. S., et al. 2004, *A&A*, 413, 415
- Kneib, J.-P., Ellis, R. S., Santos, M. R., & Richard, J. 2004, *ApJ*, 607, 697
- Lau, E. T., Kravtsov, A. V., & Nagai, D. 2009, *ApJ*, 705, 1129
- Lau, E. T., Nagai, D., Kravtsov, A. V., & Zentner, A. R. 2011, *ApJ*, 734, 93
- Lemze, D., Broadhurst, T., Rephaeli, Y., Barkana, R., & Umetsu, K. 2009, *ApJ*, 701, 1336
- Lokas, E. L., & Mamon, G. A. 2003, *MNRAS*, 343, 401
- Mantz, A., Allen, S. W., Rapetti, D., & Ebeling, H. 2010, *MNRAS*, 406, 1759
- Morandi, A., Limousin, M., Rephaeli, Y., Umetsu, K., Barkana, R., Broadhurst, T., & Dahle, H. 2011, *MNRAS*, 416, 2567

- Morandi, A., Pedersen, K., & Limousin, M. 2010, *ApJ*, 713, 491
- Piffaretti, R., Jetzer, P., Kaastra, J. S., & Tamura, T. 2005, *A&A*, 433, 101
- Piffaretti, R., & Valdarnini, R. 2008, *A&A*, 491, 71
- Planck Collaboration et al. 2011, ArXiv e-prints
- Pointecouteau, E., Arnaud, M., & Pratt, G. W. 2005, *A&A*, 435, 1
- Postman, M., et al. 2011, ArXiv e-prints
- Rapetti, D., Allen, S. W., Mantz, A., & Ebeling, H. 2010, *MNRAS*, 406, 1796
- Sarazin, C. L. 1988, X-ray emission from clusters of galaxies, ed. Sarazin, C. L.
- Sartoris, B., Borgani, S., Fedeli, C., Matarrese, S., Moscardini, L., Rosati, P., & Weller, J. 2010, *MNRAS*, 407, 2339
- Schafer, R. A. 1991, XSPEC, an x-ray spectral fitting package : version 2 of the user's guide, ed. Schafer, R. A.
- Sereni, M., Ettori, S., & Baldi, A. 2011, ArXiv e-prints
- Sereni, M., & Umetsu, K. 2011, *MNRAS*, 416, 3187
- Stark, D. P., Ellis, R. S., Richard, J., Kneib, J.-P., Smith, G. P., & Santos, M. R. 2007, *ApJ*, 663, 10
- Trotta, R. 2008, *Contemporary Physics*, 49, 71
- Vikhlinin, A., Kravtsov, A., Forman, W., Jones, C., Markevitch, M., Murray, S. S., & Van Speybroeck, L. 2006, *ApJ*, 640, 691
- Vikhlinin, A., et al. 2009a, *ApJ*, 692, 1060
- Vikhlinin, A., et al. 2009b, in ArXiv Astrophysics e-prints, Vol. 2010, astro2010: The Astronomy and Astrophysics Decadal Survey, 305–+
- Wojtak, R., Hansen, S. H., & Hjorth, J. 2011, ArXiv e-prints
- Wojtak, R., & Łokas, E. L. 2010, *MNRAS*, 408, 2442
- Yoo, J., Fitzpatrick, A. L., & Zaldarriaga, M. 2009, *Phys. Rev. D*, 80, 083514

A. APPENDIX

A.0.1. Creating artificial observations of an x-ray gas

In our analysis we have two different situations where we need to simulate a dataset. The first is as input to the MCMC routine when fitting to a given dataset. The second is where we actually simulate the dataset that has to be fitted, i.e. the mock data. The first steps for both are the same, and is described in the following; Given a set of parameterized profiles and shape we create a three dimensional x-ray gas on a grid. The local spectral information is calculated by **XSPEC**'s (see e.g. Arnaud (1996), Schafer (1991)) model mekal (<http://heasarc.nasa.gov/xanadu/XSPEC/manual/XSmodelMekal.html> and references within) at redshift zero including galactic absorption. We use five times higher spatial resolution in the inner regions compared to the outer parts, to make sure no resolution effects propagate into the results. We then project all the spectral information onto the 2D observational plane defined such that the x-gas structure is spherical symmetric in that plane. The projected data is then convolved in **XSPEC** with an instrumental response function, here chosen to be from **CHANDRA**, to create a final observed picture. In an ideal world this is the picture read out from the instrument assuming pixelation from the CCD is unimportant. In a real world, a spacial and spectral rebinning is done at this step. When we create a dataset as input to the MCMC routine, the binning is done so that it matches the binning of the observed dataset. When generating a mock dataset we do the binning such that the radial bins have the same number photon counts and the spectral bins have more than a given threshold. This ensures equally statistical weights for each bin. For the fits in this paper, we fixed the number of radial bins to 12 for all datasets. Because an x-ray gas density profile usually have a logarithmic shape, the radial bins are therefore also approximately logarithmic linear spaced. Our spectral threshold is chosen such that the number of new spectral bins are around 200, of originally 1024. This corresponds to a threshold of 20 counts per spectral bin for a $6 \cdot 10^4$ number photons observation. It was not computationally possible to scan over different binning strategies, but the chosen binning is believed to match a real case scenario fairly well.

Figure 8 (left) illustrates a noise free generated x-ray gas map with a non-spherical shape and its temperature profile. The shape and the temperature profile is the one used for 'shM2' introduced in section 3.2. The right plot in figure 8 shows two spectra generated from the region between the two black lines shown in the left plot. The spectrum in red is a free-free spectrum generated with **XSPEC** using the mean projected temperature and the spectrum in blue is the true projected spectrum, i.e. the sum of many free-free spectra each generated locally in the x-gas. The difference seen in the lower part of the right plot is basically what give us information about shape and profiles.

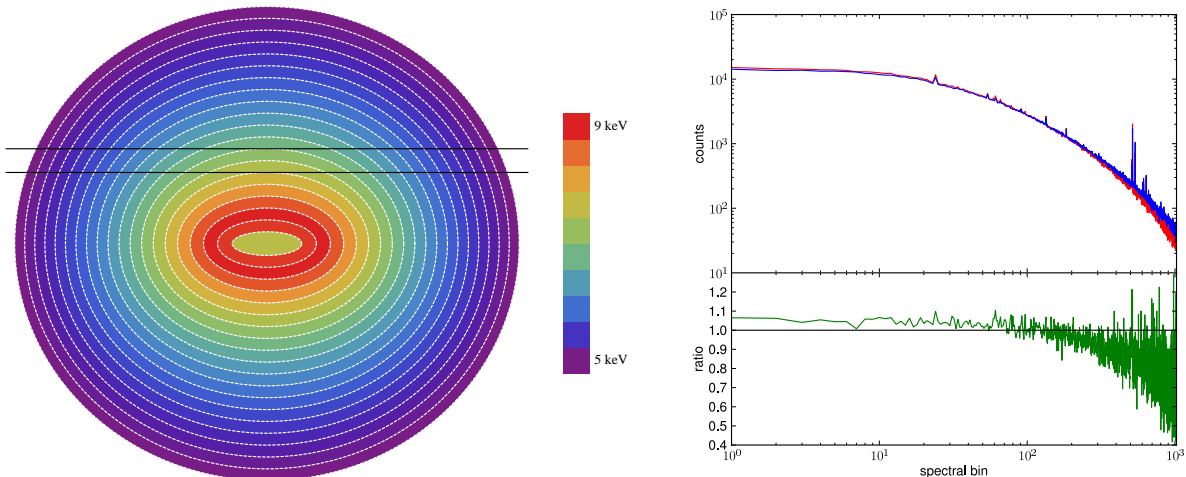


Fig. 8.— Left: A generated x-ray gas map with no noise added in the plane along the line of sight for structure 'shM2' introduced in section 3.2. The color indicates the temperature in keV. Right: Upper plot shows two spectra generated from the region between the two black lines shown in the left plot. The spectrum in red is a free-free spectrum generated with the mean x-ray temperature from the region between the two black lines in the left plot and the spectrum in blue is the true projected spectrum, i.e. the sum of many free-free spectra each generated locally in the x-gas. The lower plot shows the ratio between the red and the blue spectra.

A.0.2. Monte Carlo Technics used for this paper

We wrote a Monte Carlo Markov-Chain (MCMC) algorithm for fitting to a data set. The MCMC uses a Metropolis-Hastings sampling (Chib & Greenberg (1995)) with a flat and symmetric proposal density. The size of this proposal density was tuned to reach an acceptance rate of around 0.2-0.3 which has been shown to be the most optimal for

sampling higher dimensional distributions. The width of the proposal density along each parameter axes was tuned in units of the root mean square for the individual PDF for each parameter. For all runs the sampling space was limited by bounds on each parameter axes and realizations with a temperature profile exceeding 15 keV or going below 0.5 keV was given zero probability. Among numerous tests of possible resolution, boundary or sampling effects we tested that the codes reproduced the theoretical expected degeneracy between an overall density scaling and a fixed axis ratio along the line of sight. We tested this up to a total number of 500.000 photons. A sample of tests was also done against an independently written code which generates artificial x-ray data using the “shell binning” approach (see e.g. <http://cxc.harvard.edu/contrib/deproject/>). We tested convergence by starting chains at random places and with different scalings (number photons) of the PDF. All distributions shown in the paper are based on $5 \cdot 10^6$ samplings. The fitting results presented are based on one realization of data, marginalizing over several realizations was not computationally possible. We assumed a diagonal covariance matrix for the observed photon measurements and the noise to be gaussian.